Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 495–498, January 1992
Biochemistry

Sequence-specific DNA purification by triplex affinity capture

(triple-helix DNA/biotinylated oligonucleotide/streptavidin/magnetic separation/dinucleotide repeat)

TAKASHI ITO, CASSANDRA L. SMITH*, AND CHARLES R. CANTOR

Department of Molecular and Cell Biology, 529 Stanley Hall, University of California, and Division of Chemical Biodynamics, Lawrence Berkeley Laboratory, Berkeley, CA 94720

Contributed by Charles R. Cantor, September 27, 1991

ABSTRACT **A DNA isolation procedure was developed by using triple-helix formation and magnetic separation. In this procedure, target DNA is captured by a biotinylated oligonucleotide via intermolecular triplex formation, bound to streptavidin-coated magnetic beads, and recovered in double-stranded form by elution with a mild alkaline buffer that destabilizes the triple helix. The effectiveness of the procedure was demonstrated by a model experiment with an artificially reconstructed library and, also, by the isolation of (dT-dC)$_n$·(dG-dA)$_n$ dinucleotide repeats from a human genomic library. This procedure provides a prototype for other triplex-mediated DNA isolation technologies.**

The first triple helical structure of nucleic acids was discovered more than 30 years ago (1). While the biological roles of such structures are still open to question, their chemical characteristics have been considerably elucidated in recent works (for review, see ref. 2). The most well-characterized triplex is the one formed between a double-stranded homopurine-homopyrimidine helix and a single-stranded homopyrimidine tract. In this type of triple helix, the third homopyrimidine strand binds to the major groove, parallel to the homopurine strand of Watson–Crick double-helical DNA, via Hoogsteen hydrogen bonding. The third-strand thymine (T) recognizes adenine-thymine (A·T) base pairs forming T·A·T triplets, and the third-strand cytosine (C), protonated at its *N*-3 position, recognizes guanine-cytosine (G·C) base pairs forming C^+·G·C triplets.

Homopyrimidine oligonucleotides have been shown to form local triplexes with corresponding homopurine sites in larger double-stranded DNAs. Such oligonucleotide-directed triplex formation has been successfully applied in the recent development of sequence-specific artificial rare-cutting endonucleases (3, 4), in which oligonucleotides and equipped metal chelates or photoactive groups function as DNA binding and cleaving "domains," respectively. Also, single-site enzymatic cleavage of the yeast genome was achieved by the triplex-mediated "Achilles' heel cleavage" procedure (5), in which a triplex-forming oligonucleotide, instead of a DNA binding protein (6), was used to protect the targeted single-site from DNA methyltransferase. Such triplex-mediated DNA cleavage techniques provide valuable tools for genome analysis. Specific inhibition of DNA-binding proteins (e.g., transcription factors or replication factors) by triplex formation (7–9) may provide a principle for the development of antiviral or anticancer drugs. Based on the stability of such triplexes during gel electrophoresis (10), a unique procedure for labeling specific DNA fragments was also devised to facilitate restriction mapping of cosmid inserts (11).

To explore further applications of triple helix DNA formation, we tested the use of such structures as tools for the isolation or enrichment of specific DNA molecules from heterogeneous DNA mixtures. Such triplex-mediated procedures seem to have some unique advantages over conventional, hybridization-based methods (see *Discussion*). As a prototype for various, possible triplex-mediated DNA isolation technologies, a triplex affinity capture procedure was developed by using a biotinylated oligonucleotide and magnetic DNA separation (12) and was applied to the isolation of particular clones from plasmid libraries.

MATERIALS AND METHODS

Materials. Plasmid pTC45 (13) was a generous gift from J. S. Lee (University of Saskatchewan). An oligonucleotide BTC-20 [5′-biotinylated $(T\text{-}C)_{10}$] was synthesized by Operon Technologies (Alameda, CA). A human chromosome 21-specific plasmid library was prepared as described (14). Streptavidin-coated magnetic beads (Dynabeads Streptavidin M-280) and a magnetic particle concentrator (Dynal MPC-6) were obtained from Dynal (Great Neck, NY). Forward (24-mer) and reverse (22-mer) pUC/M13 sequencing primers were obtained from Promega. Restriction endonucleases and *Taq* DNA polymerases were purchased from New England Biolabs and Perkin–Elmer/Cetus, respectively. DNA ligation kit was obtained from Takara Biochemical (Berkeley, CA).

Triplex Affinity Capture Procedure. Plasmid DNA (≈2 μg) was incubated with 20 pmol of biotinylated oligonucleotides in 50 μl of buffer B (2.0 M NaCl/0.2 M sodium acetate/acetic acetic acid, pH 4.5–5.5) at 50°C for 2 hr. (Note that the pH values were measured at room temperature after addition of NaCl.) For reactions at higher pH (6.0 or above), sodium phosphate buffer was used instead of sodium acetate/acetic acid buffer. Streptavidin-coated magnetic beads (50 μl), washed with and resuspended in buffer B (50 μl), were added to the mixture. After further incubation for 1 hr, the beads were separated by a magnetic particle concentrator and washed eight times with 0.5 ml of the same buffer. Finally, the beads were incubated with buffer E (1.0 M Tris·HCl, pH 9/0.5 mM EDTA) for 20 min. DNA was recovered from the eluate by 1:1 (vol/vol) phenol/chloroform extraction and ethanol precipitation and was used for bacterial transformation or gel electrophoresis.

Isolation and Analysis of Dinucleotide Repeat Clones. Plasmids prepared from a chromosome 21-specific library were linearized by digestion at a unique *Hin*dIII site on each clone and subjected to the procedure described above at pH 5.5. Recovered DNAs recircularized by self-ligation using DNA concentrations of ≈0.1 μg/ml were used for transformation of competent *Escherichia coli* DH5α cells. Transformant plasmids were assayed for the presence of the repeat by a polymerase chain reaction (PCR)-based method as described (15). Thermal cycling parameters were as follows: initial denaturation at 95°C for 4 min; 24 cycles of denaturation at

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*To whom reprint requests should be addressed.

94°C at 1 min, annealing at 55°C or 58°C for 1 min, and extension at 72°C for 3 min; and a final cycle with an extension time of 10 min. Primers used were forward or reverse sequencing primers and a 27-mer oligonucleotide BamTC [CCCGGATCC(TC)$_9$]. PCR products were electrophoresed on agarose gels and detected by staining with ethidium bromide.

RESULTS

Principle of Triplex Affinity Capture. The procedure is schematically illustrated in Fig. 1. First, biotinylated homopyrimidine oligonucleotides are incubated with a DNA mixture under a mild acidic condition, which promotes protonation of cytosine and, thus, triplex formation. The triplexes formed between the oligonucleotides and target DNA molecules, containing the corresponding homopurine–homopyrimidine sequences, are subsequently bound to streptavidin-coated magnetic beads and magnetically separated from other non-triplex-DNAs. Finally, the target DNAs are recovered in double-stranded form by incubating the beads under a mild alkaline condition which destabilizes Hoogsteen hydrogen bonding (triple helix) but not Watson–Crick base pairs (double helix).

Evaluation of Basic Conditions by Using a Model System. For the evaluation of basic conditions, an artificially reconstituted plasmid library was prepared by mixing pUC19 and pTC45 (13). The latter plasmid contains a 45-base-pair (bp) run of a simple T-C repeat and forms an intermolecular triplex with exogenously added (T-C)$_n$ oligonucleotide as described (10), whereas the former does not (data not shown). A 5′-biotinylated oligonucleotide (BTC-20) was used to capture pTC45 according to the basic scheme described above. Plasmids eluted from the beads were used to transform competent *E. coli* DH5α cells. Since transformants bearing pUC19 and pTC45 form blue and white colonies, respectively, on 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-Gal)-containing agar plates, the efficiency of enrichment was calculated by simply counting the number of each colony.

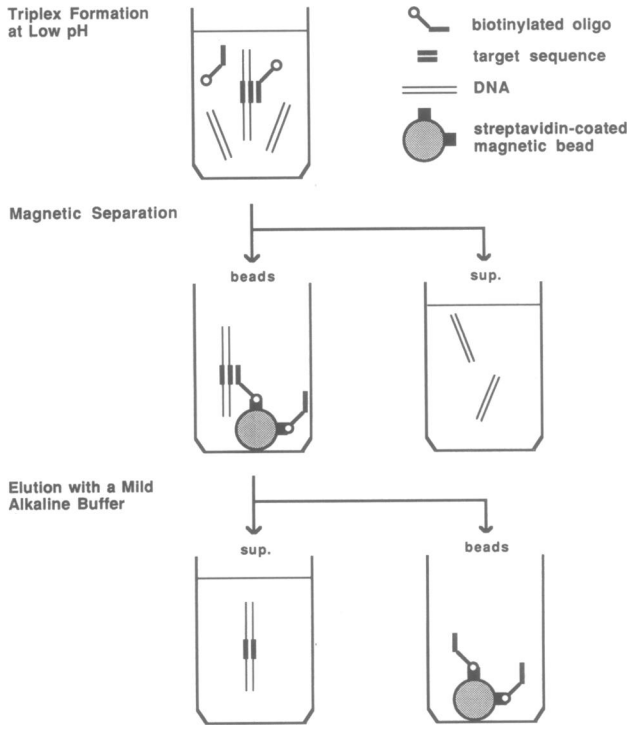

FIG. 1. Schematic illustration of triplex affinity capture procedure. See text for details.

At first, conditions were used (sodium acetate/acetic acid buffer, pH 4.8–5.0) where efficient triplex formation between pTC45 and oligo- or poly(T-C)$_n$ was observed by simple gel electrophoretic assays (refs. 10 and 16 and our unpublished observations). However, no significant enrichment of white colonies (that is, pTC45) was obtained (data not shown). It is conceivable that this is due to nonspecific interaction between DNA and streptavidin on the beads, since streptavidin, with an estimated isoelectric point of 5–6 (17), may well be positively charged under these conditions. To reduce such undesirable interactions, the ionic strength of the reaction buffer was increased by the addition of NaCl to final concentration of 2 M. [This would also increase the thermal stability of the triple helix (18).] A very efficient enrichment was obtained with the use of this modified buffer (see Table 1, for example). White colonies, that had accounted for ≈0.5% of the total colonies in the original library, were routinely enriched to account for more than 95% of the resultant library after a single round of the procedure. Occasionally, no blue colonies at all were observed. A slightly higher efficiency was reproducibly obtained when linearized plasmids were used rather than circular ones (data not shown). Thus, the use of linearized plasmids is recommended to obtain the highest efficiency purification, although this requires the additional steps of enzyme digestion before triplex formation and self-circularization before bacterial transformation.

The effects of various parameters on triplex affinity capture could be followed also by simple gel electrophoresis. For example, a mixture of linearized pTC45 and λ/*Bst*EII fragments was subjected to the triplex affinity capture procedure at various pHs, and the recovered DNAs were electrophoresed (Fig. 2). It should be noted that a 5.7-kilobase (kb) fragment derived from the left extremity of phage λ DNA, which does not contain any complete (dT-dC)$_n$·(dG-dA)$_n$ sequences longer than 7 bp but does contain some interrupted ones, also bound weakly to BTC-20 under very acidic condition (pH 4.5 or less) (see Fig. 2, lane 2). Increasing the pH up to ≈6 eliminated this weak binding without any significant reduction in recovery of pTC45 DNA (Fig. 2, lanes 3–5), whereas further increase in pH reduced the recovery considerably (Fig. 2, lanes 6 and 7). These results indicate that the stringency of the triplex affinity capture can be controlled simply by changing the pH. The same assay system also revealed the nonspecific interaction between DNA and the magnetic beads and its elimination by the high ionic strength buffer, confirming the result described above (data not shown).

Triplex Affinity Capture of (dT-dC)$_n$·(dG-dA)$_n$ Dinucleotide Repeats from a Human Genomic Library. After initial experiments with the model plasmid system, the procedure was

Table 1. Triplex-mediated enrichment of target plasmids from a reconstituted library

Colonies, no. (%)	
White (pTC45)	Blue (pUC19)
Before enrichment	
5.0×10^4	1.1×10^7
(0.5)	(99.5)
After enrichment	
4.0×10^4	0.5×10^2
(99.9)	(0.1)

Plasmids prepared from a reconstituted library were used for transformation of competent DH5α cells with or without enrichment by triplex affinity capture. pTC45 (target) and pUC19 give white and blue colonies, respectively, on 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-Gal)-containing Luria–Bertani agar plates. In this particular experiment, the enrichment was calculated to be 1.8×10^5-fold with a recovery of ≈80%.

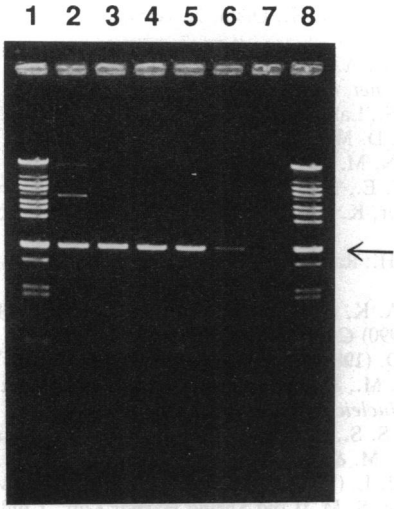

FIG. 2. The pH dependence of the triplex affinity capture reaction. A mixture of linearized pTC45 and λ/*Bst*EII was subjected to triplex-mediated DNA isolation procedure at various pHs. Recovered DNAs were electrophoresed on an agarose gel. Lanes: 1 and 8, input DNA; 2–7, reactions at pH 4.5, 5.0, 5.5, 6.0, 6.5, and 7.0, respectively. The arrow indicates the position of pTC45. At pH 4.5, the 5.7-kb *Bst*EII fragment was observed to bind weakly to BTC-20. Note that part of this fragment was recovered as a noncovalent complex, with the right arm fragment joined at the *cos* site.

applied to the isolation of $(dT\text{-}dC)_n\cdot(dG\text{-}dA)_n$ dinucleotide repeats from human genome. This sequence is a member of so-called "microsatellite" DNAs distributed throughout mammalian genomes (19, 20). It is often hypervariable in the number of the repeat units (n) from individual to individual and thus provides highly informative DNA markers for genetic linkage mapping (21–24).

Plasmids prepared from a human chromosome 21-specific library were linearized by digestion at a unique *Hin*dIII site on each clone (14) and subjected to the procedure described above. Eighteen randomly chosen clones were further analyzed. Plasmids from these clones were tested for triplex formation by binding assays similar to the one shown in Fig. 2. All these clones were able to form triplexes with BTC-20 (data not shown).

These plasmids were then tested to determine whether or not they contained $(dT\text{-}dC)_n\cdot(dG\text{-}dA)_n$ sequences using a PCR-based method originally developed by Kunkel and coworkers (15). In this assay, a $(T\text{-}C)_n$ oligonucleotide and one of the sequencing primers were combined and used as PCR primers. If the tested insert contains a $(dT\text{-}dC)_n\cdot(dG\text{-}dA)_n$ sequence(s), amplified products should be obtained by, at least, one of the primer combinations depending on the relative orientation of the dinucleotide repeat. In practice, 17 clones of the 18 tested gave distinct PCR products, part of which are shown in Fig. 3. These results demonstrated that the triplex-mediated procedure is quite effective and useful for selection of $(dT\text{-}dC)_n\cdot(dG\text{-}dA)_n$ dinucleotide repeat clones and, thus, for the subsequent development of highly informative DNA markers for genetic linkage mapping. (Further analysis of these clones will be reported elsewhere.)

DISCUSSION

A triplex-mediated DNA isolation procedure was developed using biotinylated oligonucleotides and magnetic DNA separation as schematically illustrated in Fig. 1. It proved to be quite effective both by a model experiment (Table 1) and by the selection of microsatellite DNA clones from a human genomic library (Fig. 3). It was also successfully applied to the isolation of a single-copy locus clone from a yeast

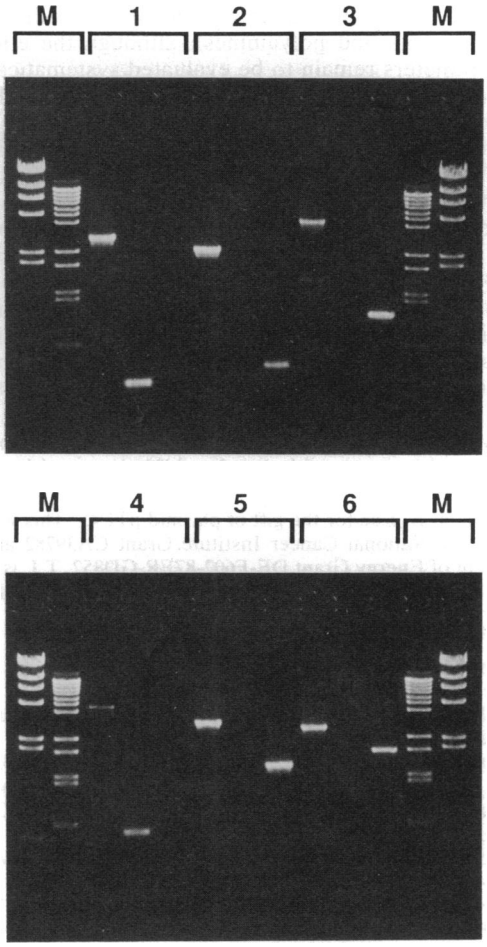

FIG. 3. PCR-based assay for the $(dT\text{-}dC)_n\cdot(dG\text{-}dA)_n$ dinucleotide repeats on six clones (nos. 1–6) purified by triplex affinity capture from a chromosome 21-specific library. PCR was performed with forward and reverse primers (each left lane), forward and BamTC primers (each middle lane), and reverse and BamTC primers (each right lane). Lanes M contain phage λ DNA digested with *Hin*dIII (outside) and *Bst*EII (inside).

genomic library (unpublished results). This triplex affinity capture procedure provides a prototype for other triplex-mediated DNA isolation technologies to be developed.

To select particular DNA clones based upon their DNA sequences, conventional hybridization-based procedures (e.g., colony or plaque hybridization) are usually employed. These well-established procedures are of great use and reliability but have some practical drawbacks. For example, they require time-consuming and labor-intensive steps of filter preparation that often limit the number of clones that can be screened. Furthermore, since these procedures include prior denaturation steps and other treatments that destroy the integrity of the target DNA molecules, one has to reisolate the corresponding clones from the original plates to obtain intact DNA molecules.

In contrast, the triplex affinity capture procedure, like RecA-mediated ones (25, 26), can be performed efficiently in the liquid phase. This eliminates the tedious filter-handling procedures and enables one to screen a larger number of clones with relative ease and, also, several libraries simultaneously in a parallel manner. It would be amenable to automation. In addition, the DNA molecules obtained are kept double-stranded, and thus readily used for various subsequent biological and biochemical manipulations. Similar to hybridization, the stringency of the triplex reaction may be controlled by tuning conditions such as pH (Fig. 2),

temperature (unpublished observations), and others (e.g., organic solvents and polyamines), although the effects of these parameters remain to be evaluated systematically.

At present, the general use of triplex-mediated DNA isolation may be somewhat restricted because of its limited target sequences (essentially, homopurine–homopyrimidine tracts). Nevertheless, judging from the success of prototype experiments described here, it is worth further pursuit. Its applicability and utility would be extended and improved by the use of some permissive mismatches in triplex formation (27, 28), alternate-strand triple-helix formation (29), other types of triple helices (30–36), including ones formed by recombinase proteins (37, 38), and artificial base analogs. A potential great advantage of the triplex-mediated approach is its applicability, in principle, to the isolation of very large DNAs, where the denaturing conditions required by almost all conventional hybridization-based approaches would inevitably lead to intolerable levels of DNA strand breakage.

We thank J. S. Lee for the gift of plasmid pTC45. This work was supported by National Cancer Institute Grant CA39782 and U.S. Department of Energy Grant DE-F602-87ER-GD852. T.I. is a recipient of a Research Fellowship from the Uehara Memorial Foundation.

1. Felsenfeld, G., Davis, D. R. & Rich, A. (1957) *J. Am. Chem. Soc.* **79,** 2023–2024.
2. Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M. & Wohlrab, F. (1988) *FASEB J.* **2,** 2939–2949.
3. Moser, H. E. & Dervan, P. B. (1987) *Science* **238,** 645–650.
4. Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J.-L., Thuong, N. T., Lhomme, J. & Hélène, C. (1987) *Nucleic Acids Res.* **15,** 7749–7760.
5. Strobel, S. A. & Dervan, P. B. (1991) *Nature (London)* **350,** 172–174.
6. Koob, M. & Szybalski, W. (1990) *Science* **250,** 271–273.
7. Maher, L. J., III, Wold, B. & Dervan, P. B. (1989) *Science* **245,** 725–730.
8. François, J.-C., Saison-Behmoara, T., Thuong, N. T. & Hélène, C. (1989) *Biochemistry* **28,** 9617–9619.
9. Hanvey, J. C., Shimizu, M. & Wells, R. D. (1990) *Nucleic Acids Res.* **18,** 157–161.
10. Lyamichev, V. I., Mirkin, S. M., Frank-Kamenetskii, M. D. & Cantor, C. R. (1988) *Nucleic Acids Res.* **16,** 2165–2178.
11. Moores, J. C. (1990) *Strategies* **3,** 23–24, 29.
12. Uhlen, M. (1989) *Nature (London)* **340,** 733–734.
13. Pulleyblank, D. E., Haniford, D. B. & Morgan, A. R. (1982) *Cell* **42,** 271–280.
14. Ito, T., Ohgusu, H., Oishi, N. & Sakaki, Y. (1991) *Genomics* **9,** 707–712.
15. Feener, C. A., Boyce, F. M. & Kunkel, L. M. (1991) *Am. J. Hum. Genet.* **48,** 621–627.
16. Lee, J. S., Latimer, L. J. P., Haug, B. L., Pulleyblank, D. E., Skinner, D. M. & Burkholder, G. D. (1989) *Gene* **82,** 191–199.
17. Green, N. M. (1990) *Methods Enzymol.* **184,** 51–67.
18. Plum, C. E., Park, Y.-W., Singleton, S. F., Dervan, P. B. & Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 9463–9467.
19. Manor, H., Rao, B. S. & Martin, G. R. (1988) *J. Mol. Evol.* **27,** 96–101.
20. Wong, A. K. C., Yee, H. A., van de Sande, J. H. & Rattner, J. B. (1990) *Chromosoma* **99,** 344–351.
21. Tautz, D. (1989) *Nucleic Acids Res.* **17,** 6463–6471.
22. Love, J. M., Knight, A. M., McAleen, M. A. & Todd, J. A. (1990) *Nucleic Acids Res.* **18,** 4123–4130.
23. Moore, S. S., Sargeant, L. L., King, T. J., Mattick, J. S., Georges, M. & Hetzel, D. J. S. (1991) *Genomics* **10,** 654–660.
24. Weber, J. L. (1990) in *Genome Analysis*, eds. Davies, K. E. & Tilghman, S. M. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), Vol. 1, pp. 159–181.
25. Honigberg, S. M., Rao, B. J. & Radding, C. M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 9586–9590.
26. Rigas, B., Welcher, A. A., Ward, D. C. & Weissman, S. M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 9591–9595.
27. Griffin, L. C. & Dervan, P. B. (1989) *Science* **245,** 967–971.
28. Belotserkovskii, B. D., Veselkov, A. G., Filippov, S. A., Dobrynin, V. N., Mirkin, S. M. & Frank-Kamenetskii, M. D. (1990) *Nucleic Acids Res.* **18,** 6621–6624.
29. Horne, D. A. & Dervan, P. B. (1990) *J. Am. Chem. Soc.* **112,** 2435–2438.
30. Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. & Hogan, M. E. (1988) *Science* **241,** 456–459.
31. Kohwi, Y. & Kohwi-Shigematsu, T. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 3781–3785.
32. Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V. & Fresco, J. R. (1988) *Biochemistry* **27,** 9108–9112.
33. Bernués, J., Beltrán, R., Casanovas, J. M. & Azorín, F. (1989) *EMBO J.* **8,** 2087–2094.
34. Beal, P. A. & Dervan, P. B. (1991) *Science* **251,** 1360–1363.
35. Pilch, D. S., Levenson, C. & Shafer, R. H. (1991) *Biochemistry* **30,** 6081–6087.
36. Orson, F. M., Thomas, D. W., McShan, W. M., Kessler, D. J. & Hogan, M. E. (1991) *Nucleic Acids Res.* **19,** 3435–3441.
37. Hsieh, P., Camerini-Otero, C. S. & Camerini-Otero, R. D. (1990) *Genes Dev.* **4,** 1951–1963.
38. Rao, B. J., Dutreix, M. & Radding, C. M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 2984–2988.